Abstract 459

TITLE: Evaluation of Non-Name-Coded Identifiers in Los Angeles County and New Jersey
AUTHORS: Marsh, K; Morgan, M; Bunch, G; Costa, S; Fleming, P; Wortley, P

BACKGROUND/OBJECTIVES: The feasibility of conducting HIV surveillance using codes in lieu of patient's names is controversial because epidemiological data are insufficient to assess the performance of codes versus names. One concern is that codes may not adequately distinguish reports on new persons and on persons previously reported to the surveillance database. We evaluated the performance of 13 codes to match $\geq$ 2 case reports representing the same person and to not match $\geq$ 2 case reports of 2 different people in Los Angeles and New Jersey. Results were compared with CDC's performance standards, which state that "#5% of incoming records should fail to match or fail not to match as appropriate in the database."
METHODS: In Los Angeles from 9/97 to 12/97 and in New Jersey from 3/98 to 1/99, information on all incoming preliminary AIDS reports, and HIV and AIDS reports, respectively, was recorded exactly as received by surveillance staff. The states used the names in these case reports to determine which reports were for new patients and which reports were duplicates. From the preliminary case reports, 13 non-name codes were also constructed. We then tested each code to determine how often it properly matched reports for the same person and did not falsely match reports for different people.
RESULTS: In Los Angeles and New Jersey, 15.0% (682/4547) and 26.9% (3,134/11,655) of the cases reported during the study periods, respectively, were new reports. In Los Angeles, performance of the codes ranged from 87.3% to 96.1% on the subset of AIDS reports in which all data used to construct the codes were complete and in which codes matched exactly. In New Jersey, during the 11-month period, performance was lower, ranging from 74.3% to 90.8%. When reports lacked some of the data needed to construct the codes, performance dropped in both study locations. Data elements required to fully complete codes were frequently missing.
CONCLUSIONS: Results from these two locations suggest that meeting CDC performance standards in non-name surveillance systems will be a challenge. Because performance is affected by characteristics of each system, such as reporting patterns and data quality and completeness, this evaluation may not predict the performance of these codes in other systems. It will be critical for states implementing surveillance based on non-name identifiers to use standardized methods to evaluate their systems to ensure comparable evaluation criteria across states and identifiers.

PRESENTER CONTACT INFORMATION

Name: Kimberly Marsh
Address: Division of Public Health Surveillance and Informatics
Epidemiology Program Office
Centers for Disease Control and Prevention
1600 Clifton Rd., NE MS C08
Atlanta, GA 30333
Telephone: (404) 639-1488
Fax: (404) 639-1546
E-mail: klm7@cdc.gov